Communicating Uncertainty in Economic Evaluations

Koffijberg, H.; de Wit, G. A.; Feenstra, T. L.

Published in:
Medical Decision Making

DOI:
10.1177/0272989X12436725

Communicating Uncertainty in Economic Evaluations: Verifying Optimal Strategies

H. Koffijberg, PhD, G. A. de Wit, PhD, T. L. Feenstra, PhD

Background. *In cost-effectiveness analysis (CEA), it is common to compare a single, new intervention with 1 or more existing interventions representing current practice ignoring other, unrelated interventions. Sectoral CEAs, in contrast, take a perspective in which the costs and effectiveness of all possible interventions within a certain disease area or health care sector are compared to maximize health in a society given resource constraints. Stochastic league tables (SLT) have been developed to represent uncertainty in sectoral CEAs but have 2 shortcomings: 1) the probabilities reflect inclusion of individual interventions and not strategies and 2) data on robustness are lacking. The authors developed an extension of SLT that addresses these shortcomings.* ***Methods.*** *Analogous to nonprobabilistic MAXIMIN decision rules, the uncertainty of the performance of strategies in sectoral CEAs may be judged with respect to worst possible outcomes, in terms of health effects obtainable within a given budget. Therefore, the authors assessed robustness of strategies likely to be optimal by performing optimization separately on all samples and on samples yielding worse than expected health benefits. The approach was tested on 2 examples, 1 with independent and 1 with correlated cost and effect data.* ***Results***. *The method was applicable to the original SLT example and to a new example and provided clear and easily interpretable results. Identification of interventions with robust performance as well as the best performing strategies was straightforward. Furthermore, the robustness of strategies was assessed with a MAXIMIN decision rule.* ***Conclusion***. *The SLT extension improves the comprehensibility and extends the usefulness of outcomes of SLT for decision makers. Its use is recommended whenever an SLT approach is considered.* ***Key words:*** *cost-effectiveness analysis; resource allocation; decision analysis; uncertainty analysis.* ***(Med Decis Making 2012;32:477–487)***

During the last 3 decades, methods for handling and presenting uncertainty in the results of economic evaluations have been developed. Most economic evaluations compare a limited set of specific and related interventions, for example, different treatment modalities for a single disease.[1] For such evaluations, the effect of uncertainty in input parameters can be assessed with probabilistic sensitivity analysis (PSA), or bootstrapping, and visualized with cost-effectiveness acceptability curves (CEAC).[2] However, if the number of interventions under consideration increases, the interpretation of the results of these methods becomes difficult. For many purposes, it has been recognized that all relevant interventions ideally should be evaluated jointly to avoid methodological inconsistencies and allow for disease dependencies as well as interactions among interventions.[3,4] Such an overall combined approach, with an explicit budget constraint, has been labeled ''sectoral'' or ''generalized'' cost-effectiveness analysis (CEA).[1,5–7] The broader perspective of sectoral CEAs is based on the notion that health care resources should be allocated across interventions and population groups such that overall population health is maximized. Only by

Received 18 January 2011 from Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands (HK, GAdW); Center for Prevention and Health Services Research, National Institute of Public Health and the Environment, Bilthoven, The Netherlands (GAdW, TLF); and Department of Epidemiology, University Medical Center Groningen, Groningen, The Netherlands (TLF). Financial support for this study was provided entirely by the Strategic Research Fund of the National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands (Grant S/260176).The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the report. The following authors are employed by the sponsor: Dr G. A. de Wit and Dr. T. L. Feenstra. The authors thank 3 anonymous reviewers, the associate editor, and the editor-in-chief for their helpful suggestions. Revision accepted for publication 17 November 2011.

Address correspondence to Hendrik Koffijberg, PhD, University Medical Center Utrecht, Julius Center for Health Sciences and Primary Care, STR 6.131, P. O. Box 85.500, 3508 GA, Utrecht, The Netherlands; telephone: +31 88 755 9713; fax: +31 88 755 5485; e-mail: h.koffijberg@umcutrecht.nl.

DOI: 10.1177/0272989X12436725

assessing all interventions simultaneously may one identify certain current interventions having unfavorable cost-effectiveness and other, not yet (fully) implemented interventions having favorable cost-effectiveness. Such identification can trigger resource reallocation from cost-ineffective to cost-effective interventions, aimed at improving population health.[1] However, despite the more complete information provided by sectoral CEAs compared with ordinary (single) CEAs, few of them have been conducted.[4] One notable exception is the ACE project, in which cost-effectiveness information was gathered in a structured way for a range of 123 different preventive interventions addressing noncommunicable diseases. Subsequently, 27 of these were evaluated comprehensively and compared with each other.[8,9] For several topics, among them alcohol abuse, obesity, and physical inactivity, optimal packages of interventions for the Australian population were estimated and efficiency frontiers depicted in a cost-effectiveness plane, showing what combinations of interventions would result in most health gains at lowest costs. Although the individual evaluations included assessment of uncertainty, the results on optimal packages were for the point estimates and only presented information on uncertainty in a qualitative way. Using a somewhat different terms, other applications evaluating a large number of interventions simultaneously were called *resource allocation* or *budget allocation* applications.[10–13] Some of these applications considered a very wide range of interventions,[12,13] whereas most were focused on a single disease or risk factor.[10,11]

We hypothesize that 2 main reasons exist for the lack of applications of sectoral CEA and resource allocation models. First, applying sectoral CEA requires extensive and complex information, for example, on the interactions between interventions and targeted diseases.[4] If decision makers are concerned with priorities over broader fields of application or concerning interventions with a large budgetary impact, however, then gathering more information and performing a sectoral CEA may be worthwhile. Second, the approach toward uncertainty in sectoral CEAs and resource allocation applications may still be improved considerably. Most published resource allocation models were deterministic and did not include an uncertainty analysis. Stochastic programming is the most straightforward extension of deterministic allocation models and has been explored in several methodological publications.[14–16] However, no applications have been published so far.

Stochastic league tables (SLTs) have been specifically developed to assess the performance of interventions in sectoral CEAs while accounting for uncertainty in costs and effects.[17,18] The SLT consists of probabilities of interventions being included in the optimal mix of interventions for a given budget, presented for a range of budgets. It has been applied to case studies in diabetes and cardiovascular disease.[19] Alternatively, multi-intervention CEACs can be used to visualize uncertainty.[20]

These approaches, however, have certain limitations. First, both the SLT and the CEAC approach focus on *interventions* and report which interventions may be part of an optimal strategy. However, this information does not allow the immediate identification of the actual optimal *strategies*, that is, the actual combination of interventions that is expected to yield the best balance between health outcomes and costs. Identification of such optimal strategies, which is paramount for decision making, requires extending the existing methods. Second, although the SLT and the CEAC approach are stochastic and account for uncertainty in the input, that is, the costs and effects of the interventions considered, they do not incorporate any uncertainty analysis of the output, that is, the probability that a specific intervention or strategy is optimal. We will illustrate these 2 limitations in 2 examples.

This article presents an extension of the SLT approach that provides information on the optimal strategies as well as the optimal interventions and incorporates a simple uncertainty analysis of the results.

METHODS

Examples

Our extension is based on the previously published method of stochastic league tables.[17,18] We applied the extension to the example in the original article by Hutubessy and others[17] as well as to a second, new example. In these examples, mutually exclusive sets of interventions are considered. Mutual exclusion in health care settings may occur when different variations of interventions, such as, for example, screening programs, are considered, of which at most 1 variation will be implemented. Interventions that can always be implemented in combination with all other interventions will constitute their own set. Hence, mutual exclusivity of interventions is not required for this analysis. Note that our examples are hypothetical and therefore do not have units of measure associated with them. Costs, resource constraints, and budgets could be expressed in, for example, US dollars, Euros, or any other currency. Effects could be expressed in terms of, for example, quality or disability adjusted life-years or some other general measure of health.

Table 1A contains the hypothetical example first outlined by Hutubessy and colleagues in Table 1 of their original paper on SLTs.[17] Three mutually exclusive sets of interventions were defined: A1-A4, B1-B3, and C1-C4. To reflect uncertainty, log-normal distributions were defined for the costs and normal distributions for the effects of each of the 11 interventions. Costs and effects were assumed to be independent for all separate interventions and between interventions, and fixed standard deviations were used for all distributions. Note, however, that the SLT method does not require these simplifications and can also be applied when data are dependent. The amount of different possible strategy combinations for this situation equals 99, as we can pick any one of or none of the interventions A1-A4, any one of or none of the interventions B1-B3, and any one of or none of the interventions C1-C4; that is, we have $5 \times 4 \times 5 = 100$ combinations but ignore the one in which no intervention is picked at all. In the remainder, combinations of interventions will be referred to as strategies. For each strategy, the costs and effects are defined as the sum of the costs and effects of all individual interventions it contains.

Table 2A contains the data for the second hypothetical example. Here, 5 mutually exclusive sets of interventions are defined: A1-A2, B1-B2, C, D1-D5, and E1-E3. To reflect uncertainty, normal distributions were defined for the costs and for the effects of each of these 13 interventions. Costs and effects were correlated within interventions ($\rho = 0.50$), and costs as well as effects were also correlated between interventions in the same mutually exclusive set ($\rho = 0.25$). In this second example, the number of possible strategies equals $(3 \times 3 \times 2 \times 6 \times 4) - 1 = 431$.

Algorithm for SLT

We reproduced the WHO-CHOICE (CHOosing Interventions that are Cost-Effective) implementation of the SLT method, called MCLeague, in the R software environment for statistical computing and graphics (v2.11.1, code available on request).[21] Our implementation was as straightforward as possible and consisted of the following steps, for a given budget X:

1. Generate all possible valid strategies, that is, all possible combinations of interventions, from the mutually exclusive sets.
2. Generate a large number of samples, for example, 25,000, of costs and effects for each separate intervention.
3. Determine the following for each sample:
   a. The total cost and total effect of each strategy by summing up the costs and summing up the effects of the interventions comprising the strategy
   b. The strategy that provides the largest health effect among all strategies with a total cost less than or equal to the given budget X. This is the optimal strategy for that sample; record this strategy and the health effect it provides in this sample
4. Determine the following based on all samples:
   a. For each strategy: how many times it was the optimal strategy
   b. For each intervention: how many times it was contained in an optimal strategy

Here, Step 3b implements a specific optimality criterion involving a tradeoff between the health effects provided by strategies and their likelihood of not exceeding the budget. In each sample, only strategies that stay within budget may be optimal. Applying a different optimality criterion may result in different optimal strategies. However, the steps of the optimization algorithm itself remain unchanged, apart from the selection of the optimal strategy in Step 3b.

Note that Step 4a is an addition to the original SLT approach, resulting in information about the optimality of strategies (combinations of interventions) rather than separate interventions.

Extending SLT with Uncertainty Analysis

We also extended the stochastic league table method with a basic form of uncertainty analysis. To this end we applied probabilistic analogies to the non-probabilistic MAXIMIN and MAXIMAX principles. The MAXIMIN principle is based on evaluating each option in terms of its *worst* possible outcome and then selecting the option with the best of these worst outcomes (maximum of the minimums). Conversely, the MAXIMAX principle is based on evaluating each option in terms of its *best* possible outcome and then selecting the option with the best of these best outcomes (maximum of the maximums). Hence, the MAXIMIN principle can be viewed as a cautious, or

Table 1A Costs and Effects of 3 Independent Sets of Mutually Exclusive Interventions in Example 1, First Proposed by Hutubessy and Colleagues[17]

Intervention	Total Costs		Total Effects	
	$\bar{x}$	s	$\bar{x}$	s
A1	120	20	1	2
A2	140	20	5.5	2
A3	170	20	3	2
A4	190	20	7	2
B1	100	20	12	2
B2	120	20	17	2
B3	150	20	20	2
C1	50	20	22	2
C2	70	20	24.5	2
C3	120	20	29	2
C4	170	20	31	2

Table 1B Corresponding Stochastic League Table (SLT) Based on Our Implementation

Intervention	Resource Availability							
	50	100	150	200	300	400	600	800
A1	0	0	0	0	2	8	1	0
A2	0	0	0	0	9	47	28	28
A3	0	0	0	0	0	4	4	4
A4	0	0	0	0	1	21	67	67
B1	0	2	37	15	3	3	0	0
B2	0	0	24	45	33	32	15	15
B3	0	0	2	40	64	64	85	85
C1	51	22	43	36	5	5	0	0
C2	13	63	24	49	10	18	0	0
C3	0	14	27	14	49	47	24	24
C4	0	0	5	1	35	31	76	76

Note: The SLT presents the probability of inclusion (%) of the intervention in the optimal strategy at different budget levels

pessimistic, approach, whereas the MAXIMAX principle can be viewed as an optimistic approach. To assess uncertainty, the variation in ''optimality'' of the considered strategies can be assessed; that is, it can be determined how robust the estimated probability for strategies of being optimal (Step 4a) actually is.

Therefore, we extended our previous implementation to assess the variation in performance of all selected strategies (i.e., those with >5% chance of being optimal as determined in Step 4a) by modifying Step 2 to generate more samples (i.e., 100,000) and adding Step 5, defined as follows:

5. Determine the following for each (selected) strategy:
   a. The subset of samples in which this strategy provides relatively small health effects (defined below)
   b. The probability that this strategy was the optimal strategy in the subset of 5a
   c. The subset of samples in which this strategy provides relatively large health effects (defined below)
   d. The probability that this strategy was the optimal strategy in the subset of 5c

In this Step 5, the performance of each strategy is assessed based on conditional sampling. For each strategy we independently defined samples with relatively small health effects (Step 5a) as all samples with health effects below the 25th percentile of the distribution of health effects for that strategy. Samples with large health effects (Step 5c) were defined as samples in which the health effects exceeded the 75th percentile of the corresponding distribution for

Table 1C Optimality of Strategies and Interventions for Example 1 and a Fixed Budget of 300

Best Strategies		Probability of Being Optimal	Cumulative Probability of Optimality	Probability of Intervention Being Optimal Due to Incorporation in Best Strategy 1–4										
				A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4
1	Interventions B3-C3	41%	41%	—	—	—	—	—	—	41%	—	—	41%	—
2	Interventions B2-C4	18%	59%	—	—	—	—	—	18%	—	—	—	—	18%
3	Interventions B3-C4	15%	74%	—	—	—	—	—	—	15%	—	—	—	15%
4	Interventions B2-C3	8%	82%	—	—	—	—	—	8%	—	—	—	8%	—
		82%		0%	0%	0%	0%	0%	26%	56%	0%	0%	49%	33%
				Probability of Intervention Being Optimal Due to Incorporation in *Any* Remaining Strategy										
				A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4
All remaining strategies		18%	100%	2%	9%	0%	1%	3%	7%	8%	5%	10%	0%	2%
Total probability of intervention being optimal				2%	9%	0%	1%	3%	33%	64%	5%	10%	49%	35%

Note: Interventions not included in the best strategies are indicated by dashes.

that strategy. Hence, Step 5 allows the assessment of the performance of strategies for situations in which they provide either (much) more or less health effects than expected, independent of the health effects provided by other strategies. Steps 5a and 5b, evaluating only samples with the worst 25% health effects per strategy, may be seen as a probabilistic equivalent of the nonprobabilistic MAXIMIN principle. Conversely, Steps 5c and 5d may be viewed as a probabilistic equivalent of the MAXIMAX principle. The larger number of samples generated in Step 2 (100,000 instead of 25,000) is required to again obtain 25,000 samples after conditioning on a selection of 25% of all generated samples in Step 5. Note that because of conditional sampling, the sample base (subset of samples) corresponding to each strategy will be different.

RESULTS

The standard presentation of the results of an SLT is a table showing the probability of inclusion of each intervention in the optimal strategy given certain levels of resource availability (Step 4b from the algorithm). Table 1B represents the results of our SLT implementation for the first example (defined in Table 1A), based on 100,000 samples. Table 2B represents the results of our SLT implementation for example 2 (defined in Table 2A), also based on 100,000 samples. A rerun with the same number of samples provided identical results for Tables 1B and 2B, indicating that sufficient samples were generated to yield robust results. Here, the SLTs in Tables 1B and 2B clearly illustrate the 2 limitations mentioned in the introduction: 1) They do not provide any explicit information on the optimality of strategies, and 2) they do not incorporate any uncertainty analysis of the output. For instance, for a budget of 400 in the first example (Table 1B), interventions A2, A4, B2, B3, C3 and C4 all look promising, but which combinations of these interventions actually make up promising strategies cannot be determined. Furthermore, the robustness of the probabilities of optimality for these interventions is not clear. When interventions perform better (or worse) than expected, this may affect their probability of being optimal to a different extent. That is, some interventions always show a high probability of being optimal, whereas others have large variability in their outcomes and thus may be excluded from the optimal set if they perform worse than expected. Estimating this variation, due to better (or worse) than expected performance, which is not shown in the standard SLT, allows assessment of the robustness of the SLT results. For example, in Table 1B, if the probability for C3 with point estimate 47% could actually range from 20% (if C3 performs very poorly) to 80% (if C3 performs very well) and if for C4, with point estimate 31%, this range could be 10% to 60%, we would be less certain that C3 is better than C4 than when these respective ranges could be 40% to 55% and 20% to 42%.

Table 2A Costs and Effects of 5 Independent Sets of Mutually Exclusive Interventions for Example 2

	Total Costs (×1000)		Total Effects	
Intervention	$\bar{x}$	s	$\bar{x}$	s
A1	212,000	84,853	9700	1162
A2	46,800	20,000	6000	721
B1	174,000	69,282	29,000	3479
B2	18,300	10,000	26,000	3114
C	127,000	50,990	57,000	6841
D1	37,000	17,321	19,000	1500
D2	−68,000	14,142	21,000	7099
D3	95,000	112,250	40,000	1500
D4	385,000	85,440	50,000	1800
D5	307,000	102,470	57,000	5604
E1	750,000	300,000	52,000	6237
E2	336,000	134,536	51,000	6124
E3	1000	400	38,000	4561

Note: Here, costs and effects are correlated within interventions ($\rho = 0.50$), and costs as well as effects are correlated between interventions in the same mutually exclusive set ($\rho = 0.25$).

Table 2B Corresponding Stochastic League Table (SLT) Based on Our Implementation

	Resource Availability (×1,000,000)							
Intervention	**50**	**100**	**200**	**300**	**400**	**600**	**800**	**1000**
A1	5	6	11	18	20	25	42	53
A2	53	44	64	60	57	61	50	41
B1	3	3	7	14	16	21	36	44
B2	85	85	92	86	84	79	64	56
C	44	78	100	100	100	100	100	100
D1	1	2	5	7	2	0	0	0
D2	78	76	58	26	10	1	0	0
D3	21	22	35	58	56	15	7	2
D4	0	0	0	0	2	5	4	5
D5	0	0	1	9	29	79	90	93
E1	1	1	1	2	2	2	4	6
E2	2	2	4	8	13	21	47	76
E3	97	97	95	90	85	77	50	18

Note: The SLT presents the probability of inclusion (%) of the intervention in the optimal strategy at different budget levels.

Related Optimality of Interventions and Strategies

From our algorithm we additionally determined the probability that strategies are optimal (Step 4a from the algorithm). Table 1C shows the results of the algorithm for the first example (defined in Table 1A) and a fixed budget of 300. Similarly, Table 2C shows the results corresponding to the second example (defined in Table 2A) and a fixed budget of 100 (×1,000,000). In Tables 1C and 2C, all strategies are shown that have a probability of being optimal of at least 5%. This 5% threshold was set to make sure that only strategies that definitely would not interest decision makers are dropped. In practice, decision makers will have their own (implicit) threshold above which strategies may be deemed eligible for implementation, and our 5% can be seen as the lowest imaginable threshold resulting in a table with all possibly relevant strategies. In Table 1C, the resulting 4 "best" strategies are B3-C3, B2-C4, B3-C4, and B2-C3 and have a respective probability of being optimal of 41%, 18%, 15%, and 8%. Combined, these strategies have a 41% + 18% + 15% + 8% = 82% probability of being optimal, indicating that the 99 − 4 = 95 remaining strategies have a combined probability of 100% − 82% = 18% of being optimal.

Table 2C Optimality of Strategies and Interventions for Example 2 and a Fixed Budget of 100 (×1,000,000)

			Probability of Intervention Being Optimal Due to Incorporation in Best Strategy 1–4												
Best Strategies	**Probability of Being Optimal**	**Cumulative Probability of Optimality**	**A1**	**A2**	**B1**	**B2**	**C**	**D1**	**D2**	**D3**	**D4**	**D5**	**E1**	**E2**	**E3**
1 Interventions B2-C-D2-E3	28%	28%	—	—	—	28%	28%	—	28%	—	—	—	—	—	28%
2 Interventions A2-B2-C-D2-E3	22%	50%	—	22%	—	22%	22%	—	22%	—	—	—	—	—	22%
3 Interventions C-D2-E3	10%	60%	—	—	—	—	10%	—	10%	—	—	—	—	—	10%
4 Interventions A2-B2-D2-E3	7%	67%	—	7%	—	7%		—	7%		—	—	—	—	7%
5 Interventions B2-C-D3-E3	6%	73%	—		—	6%	6%	—		6%	—	—	—	—	6%
6 Interventions A2-B2-C-D3-E3	5%	78%	—	5%	—	5%	5%	—	—	5%	—	—	—	—	5%
	78%		0%	34%	0%	68%	71%	0%	67%	11%	0%	0%	0%	0%	78%
			Probability of Intervention Being Optimal Due to Incorporation in *Any* Remaining Strategy												
			A1	**A2**	**B1**	**B2**	**C**	**D1**	**D2**	**D3**	**D4**	**D5**	**E1**	**E2**	**E3**
All remaining strategies	22%	100%	6%	10%	3%	17%	7%	2%	9%	11%	0%	0%	1%	2%	19%
Total probability of intervention being optimal			6%	44%	3%	85%	78%	2%	76%	22%	0%	0%	1%	2%	97%

Note: Interventions not included in the best strategies are indicated by dashes.

Subsequently, we can determine the probability that an intervention is part of 1 of the 4 best strategies and the probability that it is part of the remaining strategies. For example, intervention A1 is not contained in any of the 4 best strategies but is contained in 2% of all optimal strategies. Intervention B3, however, is contained in best strategies 1 and 3, which have a probability of being optimal of 41% and 15%, as well as in 8% of the optimal remaining strategies. Therefore, intervention B3 has a probability of 41% + 15% + 8% = 64% of being included in an optimal intervention. In Table 2C, 6 ''best'' strategies are shown with a combined probability of being optimal of 73%. Results in Table 2C can be interpreted similarly to results shown in Table 1C, even though in this second example costs and effects were correlated instead of independent. Note that in Tables 1C and 2C we distinguish between the optimality of strategies and the optimality of single interventions.

Extending SLT with Robustness Analysis

Figure 1 shows an example of the selection of samples in the additional step—Step 5. In Figure 1, the distribution of health effects is shown for 2 example strategies: A1-B2-D3 and A2-D2-E2. In addition, the distribution of the conditional samples corresponding to the 25% worst and 25% best health effects for these strategies is shown. Here, the expected costs (not visualized) and effects are 325,300 and 75,700 for strategy A1-B2-D3 and 314,800 and 78,000 for strategy A2-D2-E2, respectively. Hence, strategy A2-D2-E2 would be preferred over strategy A1-B2-D3, when only expected costs and effects are considered, as it is expected to be cheaper and more effective. However, from Figure 1 it is apparent that the health effects provided by strategy A1-B2-D3 are more certain (i.e., display less variation) than the health effects provided by strategy A2-D2-E2. As a result, the expected health benefits in the worst 25% samples of strategy A1-B2-D3 are larger than those expected for in the worst 25% samples of strategy A2-D2-E2. This indicates that strategy A1-B2-D3 is more robust than strategy A2-D2-E2 and indeed would be preferred when applying a MAXIMIN decision rule. Conversely, the latter strategy would be preferred in a MAXIMAX decision rule (best 25% samples), where it is expected to provide more health effects than the former strategy.

Table 3A presents the results from reassessing the first example (defined in Table 1A) for 3 predefined budget values: 50, 400, and 800. Similarly, Table 3B shows results for the second example (defined in Table 2A) and budget values of 50, 200, and 600 (×1,000,000). For these (arbitrarily) selected budget values, Tables 3A and 3B indicate which strategies have an estimated probability of being optimal of at least 5%. Furthermore, the robustness of these strategies is indicated, with a worst and best probability of optimality, respectively corresponding to

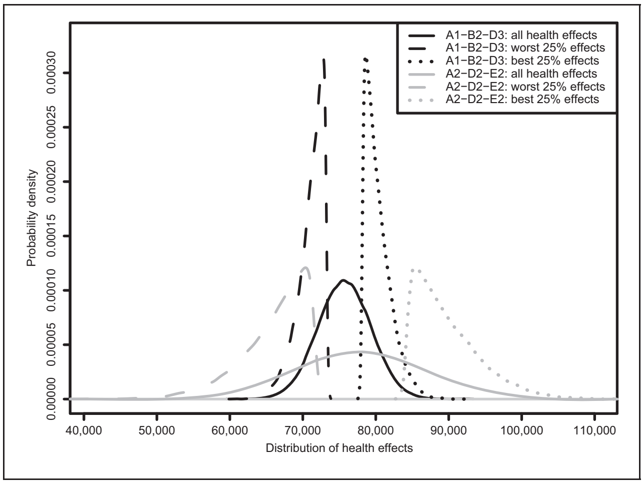

Figure 1 Distribution of health effects for 2 strategies, A1-B2-D3 and A2-D2-E2, that are part of Example 2, as defined in Table 2A. For both strategies, the unconditional distribution of health effects is visualized as well as the conditional distribution of both the 25% samples with worst and the 25% samples with best health effects. Here the expected effects are 75,700 for strategy A1-B2-D3 and 78,000 for strategy A2-D2-E2, respectively. The conditional worst/best expectations are 71,080/80,340 for strategy A1-B2-D3 and 66,080/90,000 for strategy A2-D2-E2.

the separate analysis of the 25% samples with worst and best health effects for each strategy.

Table 3A shows that in our first example, for a budget of 50 the strategy ''only C1'' is optimal and for a budget of 800 the strategy A4-B3-C4 is optimal. For a budget of 400, no single strategy has a probability of being optimal that exceeds 20%. In general, high probabilities of being optimal will only occur in situations where the options are limited, that is, when the budget is either very low, and few strategies can be afforded, or very high, and only the health effects determine the optimality of strategies.

For budgets of 50 and 800 in this first example, cautious or optimistic decision makers would be likely to reach the same conclusion as neutral decision makers. For these budget values, the optimal strategies seem to be quite robust for either best-case or worst-case scenarios. However, for a budget of 400, strategy B3-C4 has a slightly lower probability of being optimal than strategy A2-B3-C3 when mean values are considered but a slightly higher probability of being optimal when a best-case scenario is considered. None of the strategies performs very well in the worst-case scenario.

From Table 3B it is clear that for none of the selected budgets of 50, 200, and 600 (×1,000,000) does there exist a strategy with probability of being

Table 3A Results of the Uncertainty Analysis for Example 1 Defined in Table 1A, for Budget Values of 50, 400, and 800

Budget	Strategy[a]	x	Worst[a]	Best[a]
50	C1	51	49	54
	C2	13	10	14
	Total selected	64		
400	A2-B3-C3	19	6	32
	B3-C4	15	2	34
	A2-B2-C3	11	1	27
	A2-B2-C4	6	1	12
	A2-B3-C2	6	0	18
	Total selected	57		
800	A4-B3-C4	44	15	73
	A2-B3-C4	18	3	40
	A4-B3-C3	14	2	32
	A4-B2-C4	7	0	20
	A2-B3-C3	6	0	16
	Total selected	89		

Table 3B Results of the Uncertainty Analysis for Example 2 Defined in Table 2A, for Budget Values of 50, 200, and 600 (× 1,000,000)

Budget (×1,000,000)	Strategy[a]	x	Worst[a]	Best[a]
50	A2-B2-D2-E3	33	27	38
	B2-C-D2-E3	17	26	9
	C-D2-E3	10	11	8
	A2-B2-C-D2-E3	7	14	2
	B2-D3-E3	6	7	5
	A2-B2-D3-E3	6	8	5
	Total selected	79		
200	A2-B2-C-D2-E3	36	25	43
	A2-B2-C-D3-E3	17	22	11
	B2-C-D3-E3	12	13	10
	B2-C-D2-E3	7	2	13
	A1-B2-C-D2-E3	6	6	5
	Total selected	78		
600	A2-B2-C-D5-E3	32	22	38
	A1-B2-C-D5-E3	14	15	11
	A2-B1-C-D5-E3	10	11	7
	A2-B2-C-D3-E2	7	3	13
	A2-B2-C-D5-E2	5	9	1
	Total selected	68		

a. Worst performance is based on the probability of being optimal in the 25% samples providing the worst health benefits per strategy, independent of other strategies. Best performance is based on the probability of being optimal in the 25% of the samples providing the best health benefits per strategy, independent of other strategies. Due to conditional sampling, the sum of the probabilities of being optimal may exceed 100% in the Worst and Best performance columns.

optimal exceeding 36% in the expected scenario (mean) and 43% in the best-case scenario. Here the effect of correlation between costs and effects, within and between interventions, becomes apparent. For example, for a budget of 50 (×1,000,000), strategy B2-C-D2-E3 has expected probability of being optimal of 17%, which increases to 26% in the worst-case scenario and decreases to 9% in the best-case scenario. Conversely, for strategy A2-B2-D2-E3 and the same budget, this probability changes from 33% to 27% (worst case) and to 38% (best case). This difference stems from the fact that for all strategies, the probability of being optimal is limited by a combination of the health effects provided and the costs incurred. When the health effects of a strategy are relatively low, then the probability of optimality will decrease further when a worst-case scenario (with respect to health effects) is considered, and costs do not play any role. However, when the health effects of a strategy are relatively high, then the probability of optimality will largely depend on the chance to stay within budget. Given the positive correlation of 0.5 for costs and effects within interventions, considering worst-case scenarios with respect to health effects will result in the selection of samples with relatively low costs. Hence, strategies will have lower costs than expected, resulting in increased performance for expensive strategies with large health effects and decreased performance for (cheap) strategies with limited health effects. Indeed, for the strategies B2-C-D2-E3 and A2-B2-D2-E3, we find respective probabilities of 98% and 2% of exceeding the 50 (×1,000,000) budget and health effects of 149,200 (95% confidence interval 131,582–166,895) and 90,990 (95% confidence interval 73,259–108,703). Obviously, the first strategy can only be optimal when it is less expensive than expected, whereas the latter can only be optimal if it results in more health effects than expected.

To assess the effect of evaluating a subset with 25% of the best and worst samples per strategy, we repeated the analyses presented in Tables 3A and 3B with subsets of 10% and 50% of the best and worst samples. We found that the probability of optimality in worst- and best-case scenarios did not change for the vast majority of strategies. For a few strategies these probabilities increased or decreased with 1% to 4%, with larger changes occurring only for strategies with larger mean probabilities (data not shown). However, the observed changes did not actually influence the selection of strategies with a probability of optimality exceeding the 5% threshold or their ranking as presented in Tables 3A and 3B.

DISCUSSION

In this article we present an extension to the SLT approach that enables uncertainty analysis. Moreover, we aim to communicate the information of most interest to decision makers, that is, which particular strategies out of all possible strategies have a high probability of being optimal. These strategies can be considered most promising, recognizing that their actual costs and health effects are still uncertain. The implementation of our extension is straightforward and the corresponding results are easy to interpret. We feel that the addition of the uncertainty analysis is valuable, as the existing SLT approach does not explicate the overall variation in the health effects provided by interventions and strategies.

Limitations of Our Extension

With the current unoptimized implementation it may become infeasible to perform the proposed analysis when the total number of interventions considered becomes too large, as would also be the case with the standard SLT approach. When testing, we have considered up to 20 interventions and up to 20,000 strategies without running into computational problems, although time requirements increased from a few minutes to a few hours. Computational optimization may be applied so that larger sets of interventions or more strategies can be assessed. However, one may wonder whether analyses of very large sets of unrelated interventions are attractive, especially given their information requirements. In practice, a simultaneous analysis of the performance of all interventions aimed at a specific disease, for example, diabetes, is both feasible and most likely to be relevant.

Ultimately, the cost-effectiveness of interventions and of strategies is only a single aspect to be considered in priority setting in health care. Therefore, small differences between strategies are unlikely to have actual consequences for decision making, and the main benefit of the proposed approach is identification of strategies that have at least a nonnegligible probability of being optimal. This relevance check was implemented by the probability threshold of 5% we applied in the plots and tables. When even the highest probabilities of being optimal are rather low, decision makers could ask for a value of information analysis prior to selecting a strategy, to see whether postponing the decision and collecting additional data would be worthwhile.

In our examples we demonstrated that the proposed extension can be directly applied to both independent and correlated cost and effect data. Depending on costs and effects of strategies, correlation between costs and effects both within and between interventions may have substantial effect on the resulting probabilities of optimality.

For our uncertainty analysis we defined different hypothetical situations with respect to the health effects provided by the optimal strategy, for a given budget. Alternatively, we could have defined hypothetical situations based on the cost incurred by the optimal strategy, or the incremental cost-effectiveness ratio (ICER) of the optimal strategy. To keep in line with the original SLT approach, we assumed a fixed budget. Therefore, conditioning on health effects is rather intuitive, as in practice we are uncertain about the health effects that can actually be obtained within that budget. Finally, we defined different hypothetical situations in which interventions yield lower or higher health effects than expected. The size of the subset (25% of the worst or best samples) was chosen arbitrarily. Investigation of other sizes of this subset (10%, and 50%) indicated that our results and conclusions were robust with respect to subset size. This limited effect of subset size in our examples may be partially due to the normal distributions that were assumed for effects in both examples and also for costs in the second example. When less regular and smooth distributions are used for costs and effects, the influence of the size of the considered subset of samples may be larger.

Comments on the SLT Approach

After the introduction of the SLT approach, some discussions have taken place on both technical and methodological aspects of the approach.[22–25] In particular, our extension is not intended to address the initial objections to the SLT approach, warranted or not, that its use could lead to potential inefficiencies through 1) the dependence of the probabilities of inclusion on decisions related to other programs and 2) the failure to consider the opportunity costs of obtaining increased health benefits.[24,25] Rather, our extension was developed to improve the usefulness of results obtained with the SLT approach, for researchers preferring this approach. From previous discussions it can be concluded that few alternative approaches, such as confidence or credible intervals for the net benefit of strategies, described by Groot Koerkamp and others[26] and Fenwick and Briggs,[27] and the incremental benefit curve, described by Bala and others,[28] are available in addition to the SLT approach. However, it seems that neither the SLT approach nor any of its alternatives offer the single best solution to analyzing uncertainty in sectoral economic evaluations. Rather, any approach has its advantages and disadvantages. For the SLT approach, the most notable advantage is that it starts from a budget constraint, which may be the most relevant aspect of decision making from a policy perspective. However, this may also be seen as a disadvantage, since the implied risk attitude toward costs is quite high. Alternative approaches have more explicitly formulated the tradeoff between money and health benefits and the risk attitude toward uncertain outcomes on one or both of these.[14] These, however, may lack practical applicability since they require decision makers to be able to formulate an explicit objective function over risky outcomes. This complicates the comparison between results from these different types of analyses.

CONCLUSION

In settings in which sectoral CEA can be applied, care should be taken to present results from the analysis as clearly and succinctly as possible to decision makers in order to improve the usefulness of such results. Uncertainty analysis of the results also adds to that usefulness by providing decision makers with the opportunity to assess the robustness of performance for strategies. Therefore, we recommend the use of our extension whenever an SLT approach is considered by health economists conducting sectoral CEAs.

REFERENCES

1. Murray CJ, Evans DB, Acharya A, Baltussen RM. Development of WHO guidelines on generalized cost-effectiveness analysis. Health Econ. 2000;9:235–51.

2. Van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994;3:309–19.

3. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977;296:716–21.

4. Baltussen R, Brouwer W, Niessen L. Cost-effectiveness analysis for priority setting in health: penny-wise but pound-foolish. Int J Technol Assess Health Care. 2005;21:532–4.

5. Stinnett AA, Paltiel AD. Mathematical programming for the efficient allocation of health care resources. J Health Econ. 1996;15:641–53.

6. Hutubessy R, Chisholm D, Edejer TT. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. Cost Eff Resour Alloc. 2003;1:8.

7. Hutubessy RC, Baltussen RM, Torres-Edejer TT, Evans DB. Generalised cost-effectiveness analysis: an aid to decision making in health. Appl Health Econ Health Policy. 2002;1:89–95.

8. Gordon L, Graves N, Hawkes A, Eakin E. A review of the cost-effectiveness of face-to-face behavioural interventions for smoking, physical activity, diet and alcohol. Chronic Illn. 2007;3:101–29.

9. Vos T, Carter R, Barendregt J, Mihalopoulos C, Veerman L, Magnus A, et al. Assessing cost-effectiveness in prevention (ACE-prevention): final report. 2010. Available from: URL: www.sph.uq.edu.au/bodce-ace-prevention.

10. Brandeau ML, Zaric GS, Richter A. Resource allocation for control of infectious diseases in multiple independent populations: beyond cost-effectiveness analysis. J Health Econ. 2003; 22:575–98.

11. Earnshaw SR, Richter A, Sorensen SW, Hoerger TJ, Hicks KA, Engelgau M, et al. Optimal allocation of resources across four interventions for type 2 diabetes. Med Decis Making. 2002;22(5 Suppl):S80–91.

12. Epstein DM, Chalabi Z, Claxton K, Sculpher M. Efficiency, equity, and budgetary policies: informing decisions using mathematical programming. Med Decis Making. 2007;27:128–37.

13. Flessa S. Where efficiency saves lives: a linear programme for the optimal allocation of health care resources in developing countries. Health Care Manag Sci. 2000;3:249–67.

14. Al MJ, Feenstra TL, Hout BA. Optimal allocation of resources over health care programmes: dealing with decreasing marginal utility and uncertainty. Health Econ. 2005;14:655–67.

15. Chalabi Z, Epstein D, McKenna C, Claxton K. Uncertainty and value of information when allocating resources within and between healthcare programmes. Eur J Oper Res. 2008;191:529–38.

16. McKenna C, Chalabi Z, Epstein D, Claxton K. Budgetary policies and available actions: a generalisation of decision rules for allocation and research decisions. J Health Econ. 2010;29: 170–81.

17. Hutubessy RC, Baltussen RM, Evans DB, Barendregt JJ, Murray CJ. Stochastic league tables: communicating cost-effectiveness results to decision-makers. Health Econ. 2001;10:473–7.

18. Baltussen RM, Hutubessy RC, Evans DB, Murray CJ. Uncertainty in cost-effectiveness analysis: probabilistic uncertainty analysis and stochastic league tables. Int J Technol Assess Health Care. 2002;18:112–9.

19. Hutubessy RC, Niessen LW, Dijkstra RF, Casparie TF, Rutten FF. Stochastic league tables: an application to diabetes interventions in the Netherlands. Health Econ. 2005;14:445–55.

20. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ. 2001;10:779–87.

21. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2010.

22. Coyle D. Who's better not best: appropriate probabilistic uncertainty analysis. Int J Technol Assess Health Care. 2003;19:540–5.

23. Baltussen RM, Hutubessy RC, Evans DB, Murray CJ. Reply to Coyle's comments on ''uncertainty in cost-effectiveness analysis: probabilistic uncertainty analysis and stochastic league tables.'' Int J Technol Assess Health Care. 2003;19:682–4.

24. Coyle D. Determining the optimal combinations of mutually exclusive interventions: a response to Hutubessy and colleagues. Health Econ. 2003;12:159–62.

25. Baltussen MPM, Hutubessy RCW, Barendregt JJ, Evans DB, Murray CJL. Formal response to ''Determining the optimal combinations of mutually exclusive interventions: a response to Hutubessy and colleagues.'' Health Econ. 2003;12:163–4.

26. Groot Koerkamp B, Hunink MG, Stijnen T, Hammitt JK, Kuntz KM, Weinstein MC. Limitations of acceptability curves for presenting uncertainty in cost-effectiveness analysis. Med Decis Making. 2007;27:101–11.

27. Fenwick E, Briggs A. Cost-effectiveness acceptability curves in the dock: case not proven? Med Decis Making. 2007;27:93–5.

28. Bala MV, Zarkin GA, Mauskopf J. Presenting results of probabilistic sensitivity analysis: the incremental benefit curve. Health Econ. 2008;17:435–40.